Lecture 2

Examining the efficiency of estimators vis-à-vis the Cramer-Rao Lower Bound.

Hessian for $N(\mu,\sigma^2)$ model:

$$H = \begin{bmatrix} -\dfrac{N}{\sigma^2} & \dfrac{-\sum_{i=1}^{N}(x_i-\mu)}{\sigma^4} \\ \dfrac{-\sum_{i=1}^{N}(x_i-\mu)}{\sigma^4} & \dfrac{N}{2\sigma^4}-\dfrac{\sum_{i=1}^{N}(x_i-\mu)^2}{\sigma^6} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial^2 l(x;\mu,\sigma^2)}{\partial\mu^2} & \dfrac{\partial^2 l(x;\mu,\sigma^2)}{\partial\mu\partial\sigma^2} \\ \dfrac{\partial^2 l(x;\mu,\sigma^2)}{\partial\sigma^2\partial\mu} & \dfrac{\partial^2 l(x;\mu,\sigma^2)}{(\partial\sigma^2)^2} \end{bmatrix}$$

$$H_{11} = -\frac{N}{\hat{\sigma}^2} < 0$$

$$H(\hat{\mu},\hat{\sigma}^2) = \begin{bmatrix} -\dfrac{N}{\hat{\sigma}^2} & \dfrac{-\sum\limits_{i=1}^{N}(x_i-\bar{x})}{\hat{\sigma}^4} \\ \dfrac{-\sum\limits_{i=1}^{N}(x_i-\bar{x})}{\hat{\sigma}^4} & \dfrac{N}{2\hat{\sigma}^4}-\dfrac{\sum\limits_{i=1}^{N}(x_i-\bar{x})^2}{\hat{\sigma}^6} \end{bmatrix} = \begin{bmatrix} -\dfrac{N}{\hat{\sigma}^2} & 0 \\ 0 & -\dfrac{N}{2\hat{\sigma}^4} \end{bmatrix}$$

Then $det[H(\hat{\mu},\hat{\sigma}^2)] = \dfrac{N^2}{2\hat{\sigma}^2} > 0$,

that is, the leading principal minors are changing signs. So H is negative-definite.

Cramer-Rao Lower Bound (CRLB)

$$CRLB = \{-E[H]\}^{-1} = I(\mu,\sigma^2)^{-1}$$

where $I(\mu,\sigma^2)$ is the information matrix

$$E[H] = E\begin{bmatrix} -\dfrac{N}{\sigma^2} & \dfrac{-\sum\limits_{i=1}^{N}(x_i-\mu)}{\sigma^4} \\ \dfrac{-\sum\limits_{i=1}^{N}(x_i-\mu)}{\sigma^4} & \dfrac{N}{2\sigma^4}-\dfrac{\sum\limits_{i=1}^{N}(x_i-\mu)^2}{\sigma^6} \end{bmatrix} = \begin{bmatrix} -\dfrac{N}{\sigma^2} & 0 \\ 0 & -\dfrac{N}{2\sigma^4} \end{bmatrix} = -I(\mu,\sigma^2)$$

as

$$E\left[\frac{N}{2\sigma^4}-\frac{\sum_{i=1}^{N}(x_i-\mu)^2}{\sigma^6}\right]=\frac{N}{2\sigma^4}-\frac{\sum_{i=1}^{N}(x_i-\mu)^2}{\sigma^6}=\frac{N}{2\sigma^4}-\frac{N\sigma^4}{\sigma^6}=-\frac{N}{2\sigma^2}$$

Then

$$CRLB = I(\mu,\sigma^2)^{-1}=\begin{bmatrix}\frac{N}{\sigma^2} & 0\\ 0 & \frac{N}{2\sigma^2}\end{bmatrix}^{-1}=\begin{bmatrix}\frac{\sigma^2}{N} & 0\\ 0 & \frac{2\sigma^2}{N}\end{bmatrix}$$

so the CRLB for unbiased estimator of μ is $\frac{\sigma^2}{N}$ while the CRLB for unbiased estimator of σ^2 is $\frac{2\sigma^4}{N}$.

Let's look at the potential MVU efficiency for the candidate estimator of μ, namely $\overline{x}$:

1) Is $\overline{x}$ unbiased?

Yes, it is. as $E(\overline{x})=E(\frac{\sum_u x_i}{N})=\frac{1}{N}\sum_i E(x_i)=\frac{1}{N}N\mu=\mu$

2) Does $Var(\overline{x})$ achieve CRLB?

Yes, it does. as $Var(\overline{x})=\frac{\sigma^2}{N}=CRLB$

By the CR theorem, we then have that $\overline{x}$ is an MVU estimator (uniqueness is not guaranteed).

Consider $s^2=\frac{\sum_i(x_i-\overline{x})^2}{N-1}$ as a potential MVU estimator of σ^2.

Recall Theorem:

$$\frac{N\hat{\sigma}^2}{\sigma^2}=\frac{(N-1)s^2}{\sigma^2}\rightarrow\chi^2_{N-1}$$

then consider

$$E[\frac{(N-1)s^2}{\sigma^2}]=E(\chi^2)=N-1$$

we can see that

$$\frac{N-1}{\sigma^2}E(s^2)=(N-1)$$

$$E(s^2)=\frac{N-1}{N-1}*\sigma^2=\sigma^2$$

therefore s^2 is an MVU candidate for the estimation of σ^2 in the $N(\mu, \sigma^2)$ model.

What about $Var(s^2)$? Does it achieve the CRLB for the estimator of σ^2?

$$Var[\frac{(N-1)s^2}{\sigma^2}] = Var(\chi^2_{N-1}) = 2(N-1)$$

This implies

$$\frac{(N-1)^2 Var(s^2)}{(\sigma^2)^2} = 2(N-1)$$

$$\text{i.e. } Var(s^2) = \frac{2(N-1)}{(N-1)^2}\sigma^4 = \frac{2\sigma^4}{N-1} > \frac{2\sigma^4}{N} \text{ (CRLB)}$$

which unfortunately is greater than the CRLB for the estimation of σ^2

Therefore, according to the CRLB theorem, s^2 may or may not be an MVU estimator of σ^2. Therefore, the results of this exercise are inconclusive in terms of determining an MVU estimator of σ^2. However, this latter inconclusive issue can be resolved by the consideration of additional (more powerful) theorems in mathematics including the theorems:

(1) Fisher-Neyman Factorization Theorem
(2) Rao-Blackwell Theorem
(3) Lehmann-Scheffe Theorem

applied in a complete-sufficient statistics framework.

The advantage of the complete-sufficient statistic approach ala the above theorems is that, not only do we obtain conclusive outcomes of efficiency comparisons but we also guarantee uniqueness of given MVU estimators. Using the complete-sufficient statistics approach to examining estimator efficiency, it can be shown that, not only is s^2 an MVU estimator of σ^2 but it is also **<u>the</u>** MVU estimator of σ^2 (i.e. is the unique MVU estimatororof σ^2).

Here is a quote taken from Fomby, Hill, Johnson <u>Advanced Econometric Methods</u> (1984, p. 36):

"The advantage of the Cramer-Rao approach is that it is easier to comprehend and apply for the novice in mathematical statistics. The Cramer-Rao approach requires regularity and the application of one theorem, the Cramer-Rao theorem. In contrast, the complete-sufficient statitics approach requires the development of two definitions, sufficiency and completenessw, and three theorems; the Fisher-Neyman factorization theorem, the Rao-Blackwell theorem, and the Lehmann-Scheffe theorem. Though intricate, the complete-sufficient statistics approach provides more definitive results. Using this approach, an unbiased estimator can be shown to be Minimum Variance Unbiased efficient or not. There is no indeterminate outcome. Moreover, once it is established that an unbiased estimator is minimum unbiased efficient by this approach, uniqueness can also be claimed. That is, the complete-sufficient statistics approach allows the determination of **<u>the</u>** minimum variance unbiased estimator."